

## Competition and Coordination in the 1960s U.S. Data Center Boom

### *Introduction*

The value of social scientific data centers, so firmly established in the academic social sciences today, was far from immediately obvious when the first ones were created in the U.S. in the mid-20<sup>th</sup> century. But by the mid-1960s, many new data centers were springing up in a climate of great anticipation and excitement. In what now seems to foreshadow the current academic buzz about “big data,” the people who devoted their professional lives to building the first data centers predicted they would fundamentally change how social science was done. And they were right.

For as important to contemporary social science as data centers have turned about to be, however, there has been little historical investigation of them as institutions. Though they are mentioned in some accounts of the social sciences during the Cold War or in assessments of the increase in quantitative proficiency among social scientists, data centers have become nearly as essential but also as invisible to contemporary social scientists as the power plants that feed electricity into their home and workplace computers. In this paper, then, I take these key institutions as my starting point, examining the early institutional lives of social scientific data centers in the U.S. and the boom climate in which they emerged.

Booms are accompanied and, in part, sustained by widespread enthusiasms. And there were plenty of people who became enthusiastic about creating social scientific data centers in the postwar U.S. In the academic world, political scientists were but one disciplinary contingent in a much larger group of academic boosters that included economists, sociologists, social psychologists, demographers, and scholars of communications. Nor was this enthusiasm confined to the academic world: a number of commercial pollsters, government officials and private and public funders also participated in it. In this paper, I am more interested in the institutional and intellectual aspirations that drove this boom rather than in investigating its origins. When this enthusiasm was at its highest pitch, why was there such intense competition between two of the largest new academic data centers in the U.S.? And how did supporters

of data centers express their hopes for these new institutions and the advances in knowledge they promised?

In much of my recent work, I have sought to understand the relation between the broader intellectual, social, political and economic features of the mid-20<sup>th</sup> century U.S. and some specific academic knowledge cultures and institutions that emerged and flourished then. In this paper, I sketch how the emergence of big, national academic data centers in the 1950s and 60s might be understood in relation to these broader conditions. I rely principally on the unpublished institutional records of one academic data center, the Inter-University Consortium of Political and Social Research (ICPSR),<sup>1</sup> the personal papers of its founder, Warren Miller, as well as a range of published works. All of these sources speak to the intellectual and professional aspirations that drove the creation of data centers as well as the most fundamental political and financial constraints they faced.

### *A Brief Pre-History of the Data Center Boom*

Academic data services vary widely between countries, as one of their earliest advocates, Stein Rokkan, observed over forty years ago. To understand why, Rokkan considered the varied history of institutions in different countries that managed social information (Rokkan 1976, 443). How readily did older, more established institutions like government bureaus and libraries adapt to computerized data-generation and analysis? And how responsive were these institutions to new demands for computer-readable data? In places where older institutions adopted computer technology quickly and even allowed social scientists some part in designing its use and application (as in Norway and Sweden), demand for non-governmental academic data centers was weak (1976, 446-448). Rokkan also noted that in countries where voluminous survey data was being produced, “pressure to institutionalize access” to that data would be strong. This, he thought, was especially true in the U.S. where private firms and academic

---

<sup>1</sup> The institution was called the Inter-University Consortium for Political Research (ICPR) at its founding in 1962. In 1975, its name was changed to the Inter-University Consortium for Political and Social Research (ICPSR). Since I am mostly concerned with its early years in the bulk of the paper, I call it by its original name.

centers were the principal producers of survey data – and where there was no central government bureau devoted to survey research. Hence, in the U.S., Rokkan concluded, there were exceptionally powerful incentives to build non-governmental data centers (1976, 448).

At the end of World War II, the U.S. Congress moved quickly to cut all financial support for the few small survey research operations that had been part of the wartime federal government. By 1946, all offices that had done surveys of U.S. citizens had wound down and closed up (Converse 1987, 180-182). Though some social scientists expected the federal government would eventually create its own survey research agency, this never happened. Instead, public opinion surveys continued to be done either by private firms (like Crossley, Gallup, Harris and Roper) and a few academic centers doing contract work for government agencies and corporations as well as some of their own grant-funded research.

How might the survey data produced by these private firms and academic centers be stored, maintained and distributed? These were the central challenges the new data centers that emerged in the postwar period were created to meet. One, the Roper Center, was closely connected to private polling firms. It stored survey data produced by the many offices of the Roper firm but also had an agreement to receive data produced by Gallup. The other, ICPR, grew out of the Survey Research Center (SRC), part of the Institute for Social Research housed at the University of Michigan. The Roper Center began functioning as a survey data archive in 1957 (Hastings 1963). ICPR formed a few years later in 1962.<sup>2</sup> As of 2017, both are still in existence; their respective websites stoutly proclaim the good health of each.

Today, neither organization needs to justify its existence by disparaging the other; each is well-established enough that it can afford to politely ignore the other. But it was not always so. In ICPR's first decade, those who ran it and even some who participated in it in more limited ways lambasted the

---

<sup>2</sup> The two entities on which I focus were not the only postwar data centers that tried to build a national, principally academic membership base. Chicago's National Opinion Research Center (NORC) built such a base and made some moves towards acquiring data beyond that produced by its researchers. Peter Rossi, NORC's director during the 1960s, also tried without success to get federal support to build a National Behavioral Data Center in Chicago. Yale's Human Relations Area Files (HRAF) was one of the first substantial social scientific data archives. Its data was gleaned primarily from anthropological and sociological research, however, not from public opinion surveys (Scheuch 2003, 387).

Roper Center and thought hard about how to “best” it. Why? To be sure, they believed “their” organization would do a better job as the go-to academic data center in the U.S. than Roper. But beyond this, I suspect some features of the boom climate in which the Roper Center and ICPR formed encouraged and intensified the bitter competition between them. And since each sought to collect, maintain and distribute a wide range of similar types of data, this proximity probably also exacerbated tensions between them.

One important caveat: in the analysis I offer below, the unpublished material I draw on all comes from ICPSR’s records or the personal papers of those closely associated with it. Therefore, though I have a good sense of how ICPR understood its position, I do not yet have a comparable view of how those affiliated with the Roper Center understood theirs, nor how the Ford Foundation or the National Science Foundation (NSF), the funders most important to their efforts, understood the merits of what each entity did or planned to do.

Social science data archives have become so familiar today that it is difficult to recover a sense of the creative work that went into their construction. One begins to get a sense of this, however, from their builders’ struggle to name this new thing. Groping for the right model, they ask: will it be a library? a repository? a bank? a center? an archive? Or something that isn’t fully like any of these? This serial trying on and setting aside of various available models is particularly noticeable in the 1950s and early 60s.

One of the earliest proposals to build such an institution, written under the auspices of a Ford Foundation grant to Columbia University’s School of Library Services, oscillates between calling what is being proposed a “library” and an “archive.” The first of the two main authors of the proposal, Columbia’s York Lucci, referred to what he envisioned as a library of survey research data (Lucci and Rokkan, 1957).<sup>3</sup> Some of the early challenges these institutions faced were indeed library-like. Not only

---

<sup>3</sup> Warren Miller would later say that it was this Lucci and Rokkan report that inspired him to build ICPR – and that ICPR turned out to be more like a library than an archive (Miller 1997).

did they have to decide who could use their materials and how; they also had to work out an analogue to an interlibrary loan system for dealing with one another.<sup>4</sup> But Rokkan, the report's second author, favored the term "archive" for what he had in mind. Like archives, he argued these entities would also have to acquire, maintain and conserve material that was not available elsewhere – and, ideally, provide "services" to the researchers who used them (1957, 118). For yet others, these imagined entities also resembled banks, in that they tried to increase the value of the data deposited in them by loaning it out for secondary analyses and then capitalizing on the new questions and conclusions it yielded in others' hands.<sup>5</sup> The possibilities for conflict between emerging organizations trying to do all of these things multiplied with the addition of each new model.

### *The Data Center Boom*

In 1963, Philip Converse of Michigan's Survey Research Center and ICPR commented that the NSF was "getting inundated by specialized repository requests" (Converse 1963c). A few years later, ICPR's technical services director Ralph Bisco published an annotated list of data archives of interest to political scientists, including six "general purpose service" archives and six more "local service" archives in the U.S. (Bisco 1966). Excepting the two federal agencies on Bisco's list,<sup>6</sup> all the others were academic archives – and all were new. The oldest had begun their work as "service archives" in the late 1950s; a number were still in the process of forming when Bisco's piece appeared in 1966. Nor was this data archive creation boom confined to the academic world. In 1965, a prestigious committee of

---

<sup>4</sup> See the discussion of the "Farmington plan" of inter-library lending in Converse (1963a, 8-9).

<sup>5</sup> When ICPR was just beginning its operations, Converse imagined it becoming an "information treasury" that would grow as research results "feed back into" it (1962, 4). Lucci (1957, 48-82) argued that the success of a library of survey research data hinged on persuading social scientists that secondary analyses of the data it stores are worthwhile undertakings. Somewhat later, ICPR's Richard Hofferbert (n.d.) expressed the hope that continued sharing and dissemination of data would lead to "a significant multiplication of return on...investments" made both by those who created it and also by those who archived and distributed it. Internal references suggest this last piece was written in the 1970s.

<sup>6</sup> Bisco includes the U.S. Bureau of Labor Statistics and the Bureau of the Census among the six U.S. general purpose service archives (1966, 99-100).

academics and government officials recommended establishing a National Data Center that would pool, standardize and then distribute data collected by different federal government agencies (Kraus 2011, 4-8; A. Miller 1971, 54-67). Fast-paced and broad in scope, the expanding data center boom was the steady backbeat of social scientific research in the 1960s.

Advances in computing technology no doubt fueled this boom. The U.S. Census Bureau had begun to experiment with storing and analyzing data on punched cards in the late 19<sup>th</sup> century (Garfinkel 2000, 17-18); but in the academic world, this approach was still a relative novelty in the 1940s and 50s. Even those academics who stored their data on punched cards did not necessarily keep them or assume they would be of any value to others. The cards were bulky but also fragile, prone to being torn by the machines that read them – machines to which many social scientists did not have reliable access in the first place. What is more, in the first fifteen years after WW II, few social scientists accorded secondary analyses of data the same level of scholarly legitimacy as studies based on authors’ original data collections.<sup>7</sup> For any and all of these reasons, a significant quantity of data stored on punched cards was discarded or destroyed up through the mid-20<sup>th</sup> century (Lucci 1957, 14-17).<sup>8</sup>

When magnetic tape began to replace the bulkier, heavier stacks of punched cards in the 1960s, storing, analyzing and distributing machine-readable data became easier. And large mainframe computers were becoming part of the research infrastructure of more universities. Both these

---

<sup>7</sup> York Lucci devotes a considerable portion of his 1957 proposal for a “library” of survey data to trying to dispel doubts about the scholarly legitimacy of secondary analysis (using an existing body of data to answer new questions). After making a number of general arguments for the value of secondary analysis (48-54), Lucci then offers several lists of works that rely on it (66-68; 134-138). The political scientist David Truman, the first curator of the Roper Collection, notes that few faculty at Williams College (where the collection was first housed) or elsewhere showed much interest in it in the late 1940s (Truman 1991, 143-145). And Erwin Scheuch, the developer of one of the first data archives in Germany, characterized what it took to get mid-20<sup>th</sup> century social scientists to take secondary analyses seriously as partially “public relations” (2003, 393).

<sup>8</sup> Lucci itemizes the neglect and destruction of data stored on punched cards in a variety of settings (14-17); he notes that in the U.S., the survey data collected during World War II by the Office of War Information gathered dust after that office closed. When NORC took steps to acquire it, so much was missing or incomplete that it was ultimately all destroyed in 1955 (17). The data that formed the basis of the *American Soldier* project was also nearly destroyed.

developments made sharing data relatively easy – no longer did “sharing” require mailing someone a large stack of punched cards or expecting them to travel to an institution with a computer capable of reading them. Technological advances like these undeniably made the data center boom possible and accelerated it once it began.

What interests me, however, is the intense competition among several data centers in the U.S. – the Roper Center and the Inter-University Consortium for Political Research (ICPR) – during this boom. Competition was not a feature of the data center boom as a whole; indeed, many subject-specific or regionally-focused data archives that sprouted up at universities across the U.S. did not compete with one another at all. In these cases, their creators identified and then occupied particular data-collecting and distributing niches, launching projects that may well have gone undone had they not taken it upon themselves to do them.<sup>9</sup> But Roper and ICPR’s ambitions were broader. For one, each sought to acquire data from a wide range of researchers; as a matter of policy, ICPR accepted just about any data that was offered to it.<sup>10</sup> Each also offered institutions and individuals access to the data they stored as well as to a range of “data services” for an annual membership fee. Finally, each (at least for a time) offered specialized data analysis training to members.<sup>11</sup>

So why was there such intense, even bitter competition between these entities? Was there something about the data they housed or the services they were offering that pulled each towards seeking a national monopoly, rather than allowing them to cooperate or just tolerate or ignore each other? To what degree was the intense competition between them encouraged by external factors – factors like the preferences of the NSF, the principal funder from whom both received support? And to what degree was

---

<sup>9</sup> Two examples of subject-specific or regional archives without competitors: Purdue Opinion Panel’s collection of its surveys of nation-wide samples of U.S. high school students (Lucci 1957, 30-31) and the proposal to the NSF in 1966 by two U.S. sociologists to build an African Social Science Data Archive (Miller 1966).

<sup>10</sup> See the discussion of this issue on p. 16 below.

<sup>11</sup> The Roper Center also offered some training courses for a time, according to Kathleen Weldon, Director of Data Operations and Communications. Personal communication, September 2015.

each patterning what they did after the well-established centralized or monopolistic data storing habits of federal agencies or private corporations?

At this point, I can offer only incomplete answers to these questions. And any such answers I offer are based on what I summarize below about how ICPR competed with the Roper Center – and how those affiliated with ICPR strategized about how to do this the most effectively.

### *Competition between ICPR and the Roper Center*

The people who made the first social science data archives in the U.S. in the 1950s and 1960s energetically affirmed high-minded values like cooperation, sharing and free exchange. On the flipside of these affirmations, however, they also criticized those amongst them who were not playing by the emerging rules – or worried that they themselves might be so criticized.

In the early proposal for a library of survey research data cited above, Lucci sought to demonstrate how ready survey researchers were for such an institution by ticking off all the social science research centers claiming they made their data freely available to any fellow researcher who asked for it. He then noted pointedly that one institution – Michigan’s Survey Research Center (SRC) – said it did not do so. To explain its policy of less than free exchange, the representative of the SRC Lucci interviewed said SRC preferred to have more control over how its data were used than simply making it available to anyone who asked would allow (Lucci and Rokkan 1957, 27-28). That SRC, one of the biggest academic survey research centers, was wary of granting others free access to its data was a significant potential obstacle to the library Lucci envisioned.

There is ample evidence of such wariness in the correspondence of key SRC figures, though it is sometimes accompanied by the worry that SRC may end up being perceived as uncooperative or miserly by others. In one case that I discuss at greater length elsewhere (Hauptmann 2017), Michigan’s Miller faced a dilemma when a prominent colleague at another institution requested a substantial chunk of SRC data. On the one hand, Miller didn’t want to give this colleague all he’d asked for, since “our apparent willingness to ‘give away’ data destroys a minor part of the prize which we were thinking of offering for

participation in the consortium” (Miller Papers 1959, 2).<sup>12</sup> But on the other hand, given that his plans for ICPR hinged upon people and institutions outside Michigan joining it, Miller saw that it would be wise to “enhanc[e] the popular image of us as a generous and cooperative institution,” particularly since he had heard some grumbles from political scientists outside Michigan about ISR’s “non-participation in the Roper Collection at Williams.” This 1959 letter of Miller’s is perhaps the earliest evidence of competition between what would become ICPR and the Roper Center. Even before ICPR was officially launched, its founder-to-be was already thinking about how to attract potential members but also how to avoid being perceived as uncooperative in comparison to the Roper Center.

Just a few years after its 1962 founding, Miller and the board members of ICPR argued amongst themselves about what their stance towards the Roper Center ought to be. For his part, as ICPR’s first director, Miller favored formally granting the Roper Center a preeminent position as “the central storage facility for survey materials collected in America as well as in other countries” (Miller 1964b, 1-2) but at the same time also continuing informal negotiations to persuade Roper to lower barriers to its data for ICPR and its members (Miller 1964a, 3-4). But others who served on ICPR’s governing council or as official representatives of member institutions lobbied hard for a more confrontational stance. For instance, John Wahlke of SUNY-Buffalo, who served on ICPR’s governing council, drafted a letter to all official representatives of ICPR’s member institutions that pushed back strongly against the overly conciliatory tone of Miller’s proposals. ICPR’s council, Wahlke wrote, “felt [Miller’s] ‘proposed principles’ tended mainly towards a framework for dividing up the world of data into spheres of influence within which each imperialist archive controls distribution of its own data and protects it against outsiders” and conceded far too many “sovereign rights for data-supplying archives” (2). Wahlke also complained that Miller’s proposals “mis-labeled” Roper (which held mainly commercial survey data) a “general” archive, especially since Wahlke projected ICPR would ultimately acquire “a far broader range of data” (Wahlke 1964a, 3a). Others affiliated with ICPR also urged that it take a more confrontational

---

<sup>12</sup> Miller often referred to ICPR as “the consortium.” All subsequent quotations from this document in this paragraph appear on p. 2.

stance towards Roper. In a letter to Miller, Robert Lane of Yale advised that ICPR just start accumulating as much “poll material” as possible, especially international material, and do so in a way that made clear that ICPR hasn’t processed it itself. This, Lane says, is the best way to break Roper’s “relatively monopolistic position,” from which it gets to “establish the rules.” If there were multiple centers that had significant amounts of poll data, Lane argues, Roper would be forced “to cooperate to stay in business.” Lane concludes that pursuing an aggressive acquisitions strategy is ICPR’s best bet for improving its position in relation to Roper: “Agree to nothing that would limit the Consortium acquisitions...and quietly move in the direction of making Michigan a general repository for a wider range of public affairs survey materials” (Lane 1964, 1).

As insistently as both Wahlke and Lane tried to persuade Miller that confronting or trying to undercut the Roper Center was preferable to a conciliatory approach, their advice seems mild compared to Theodore Lowi’s bellicose letter to Miller and Wahlke. Using a nuclear “first strike” simile, Lowi argues that though ICPR’s position was strong at the moment, “like the American Atomic monopoly, such advantages wither if not used at the proper time – which is before there are other club members.” Lowi warned that though Roper “is weak at the moment” it probably wouldn’t stay so for long. The time was right, therefore, for ICPR to move in direction of centralization: “if we aim now toward destruction of ‘state sovereignty’ we’ll have done ourselves a big favor.” Such “total fusion” is not only “desirable,” Lowi argues, but “is also the only justifiable position, because foundation and government moneys are making all these efforts possible” (Lowi 1964, 1-2; emphasis in original).

I have cited a number of passages from these 1964 exchanges not only because they make clear the intensity of ICPR’s competition with Roper but also because I find the recurrent use of state language in them intriguing. Is it simply because these people were all political scientists that they compared data archives to states so readily? Or, in addition to their own connection to political science, were external imperatives encouraging them to think about ICPR in territorial or sovereign terms? Or, finally, were they aiming to collect and control data in ways that states did in other contexts? However one might account for the introduction of this language into deliberations about the shape of academic data archives, such

state language became an important, enduring conceit in how the people who built ICPR envisioned its future.

### *Coordination in the Midst of Competition*

In 1962, the same year ICPR was launched, it also oversaw the creation of another body, the Council of Social Science Data Archives (CSSDA). This council aimed to set general policies for all member archives and to manage relations between them (Bisco 1966, 104-106; 1967, 64-66; Farber and Beck 1971, 238-239); ICPR and Roper's early competition was mediated by the policies set by this new but short-lived organization. By the end of 1969, the CSSDA was no more (Farber and Beck, 239).<sup>13</sup>

Still, during its short organizational life, the CSSDA was able to do two crucial things: first, it worked quickly to develop a cooperative federation among the newer data archives with ICPR at its center (Converse 1962b).<sup>14</sup> Then several years later, with the strength of these numbers, the group presented its ideas for "inter-archive" cooperation to Roper in 1964. It was around this time that a representative for the Roper Center, M.I.T.'s Ithiel de Sola Pool, first joined the governing board of the Council (Miller 1964a). Though the Roper Center seems to have had some representation on the Council from 1964 through 1969, in its first few years, CSSDA focused on pooling the efforts of non-Roper data repositories to make a collective pitch for NSF support. Michigan's Converse reported to the group in 1963 that he thought the time was right to make a proposal to the NSF for several reasons. For one, Henry Riecken of NSF, with whom Converse had met, seemed "downright enthusiastic" about data repositories. Converse also reported that Riecken suggested an application from CSSDA would not undercut individual entities' chances of securing NSF support (Converse 1963c). A little later, Converse

---

<sup>13</sup> For a discussion of a more robust and longer-lasting set of institutions to manage cooperation among principally European social science data archives, see Scheuch (2003, 386-392). Scheuch also comments here on what he saw as the challenges Roper's ambitions and policies posed to inter-archive cooperation in Europe.

<sup>14</sup> Among those included in the earliest CSSDA discussions were representatives from UCLA, U.C. Berkeley, the University of Chicago, Yale, the University of North Carolina and the University of Cologne (Converse 1962a; Bisco 1966, 104-105).

reported to CSSDA's coordinating committee that NSF had recently rejected a proposal from the Roper Center. Converse surmised that NSF was unhappy enough with Roper to be receptive to a proposal from CSSDA (Converse 1963c). When CSSDA began its official life as an organization in 1965 (with a constitution, by-laws and an office in New York), it did so with the NSF support Converse had predicted it was well-positioned to secure (Bisco 1966, 105; Farber and Beck 1971, 238).

NSF was clearly committed to investing significant resources into numerous social science data repositories; according to Bisco, in addition to the CSSDA, three out of four non-governmental general purpose archives received NSF support (1966, 96-99). But why? At this point, I do not have enough of a sense of NSF's policies and priorities for the social sciences in the 1960s to answer this question fully. Whatever their objectives, the grants NSF made encouraged the formation and growth of multiple social science data archives. And to the extent that its grants strengthened ICPR and made a "cooperative" venture like CSSDA possible, they also intensified the competition between ICPR/CSSDA and Roper. Though I do not at this point have any insight into NSF's policies in this area, it does at least seem possible to infer that its representatives were not actively *discouraging* competition between social science data archives. To put it another way: there are no suggestions in the ICPR records I have seen that anyone affiliated with it or with CSSDA worried that competing too aggressively with Roper would damage their chances for continued NSF support.

### *The Competitor that Wasn't: The Proposed National Data Center*

One huge uncertainty hanging over all academic data centers during the 1960s was whether the federal government would build a data center of its own – and if it did so, how its operation might impinge upon their plans. Momentum for a National or Federal Data Center had been building as early as the mid-1950s, reaching its peak a decade later. A committee of the Social Science Research Council, led by Yale economist Richard Ruggles, made one of the first public cases for a National Data Center in 1965 – a case supported by the leaders several federal agencies, especially by Raymond Bowman of the Bureau of the Budget (Kraus 2011, 6-9). For academics seeking access to federal agency data, the appeal was

clear. Not only would the proposed National Data Center handle all requests for federal data; it would also standardize data from different agencies, making comparing and combining them feasible in the foreseeable future (Kraus 2011, 9-10). For federal agencies, the National Data Center promised efficiencies of several kinds: savings of time and money once each agency was no longer required to collect and maintain its own data; and quicker, more comprehensive internal policy analyses using standardized federal data (Kraus 2011, 12).

The defenders of the proposed National Data Center, however, said little about what it offered the general public – apart from a tepid afterthought that some of its benefits to social scientists and federal policy-makers might somehow eventually benefit everyone else. The envisioned centralization of federal data, so benign, sensible and efficient to its defenders, seemed deeply threatening to its numerous public and Congressional critics. Imagining a centralized store of federal data being used for blackmail, surveillance and political persecution, its opponents made passionate cases against it that frightened many and caught supporters off-guard (Kraus 2011, 13-22). Though they attempted to improve the public image of the proposed data center and quell concerns about the potential dangers it posed, salvaging even a more modest version of this proposal proved impossible. By 1970, efforts to build a National Data Center had come to a halt (Kraus 2011, 26-30).

What consequences did the development of this project as well as its eventual collapse have for ICPR? The proposed National Data Center was unlikely to be as close a rival to ICPR as the Roper Center for several reasons. First, its proposed mission centered on managing data generated by federal agencies, rather than the survey and polling data at the heart of ICPR and Roper's holdings. Second, the National Data Center was still on the drawing board, subject to a far higher level of public scrutiny than university-based data centers. For these reasons, I think people connected with ICPR watched the unveiling and eventual collapse of the proposed National Data Center with interest but without alarm. For example, in early 1967, ICPR's Technical Services Director, Ralph Bisco, commented favorably on what the planned National Data Center might do to improve the preservation and usability of federal government data. Though he implied that the federal government was lagging a bit behind university-

based data centers in managing, standardizing and providing access to the data its many agencies generated, Bisco did not portray the National Data Center as an encroachment upon other data centers' acquisitions or services (Bisco 1967, 68-70). Instead, Bisco seemed only mildly worried that the "adverse publicity" the proposal had received from "sensational and distorted" media coverage might undermine it (Bisco 1967, 70). On balance, Bisco hoped the National Data Center proposal would survive these public trials; but ICPR and other university-based data centers faced no comparable challenges. Their continued development and success did not hinge upon the establishment of a National Data Center. Perhaps the collapse of this proposal left the field a bit more open for ICPR to acquire even broader swathes of federal government data.

#### *ICPR's State-Like Acquisitions*

In its first decade, ICPR pursued an aggressive and wide-ranging data acquisitions policy, collecting historical data on political processes as well as survey research and polling data. Some collections ICPR acquired were truly gargantuan, like the eight-ton Congressional Vote Analysis collection received from Columbia University in 1967 on a long-term loan. This collection, originally done "under WPA auspices" in the 1930s, recorded and analyzed all Congressional votes from the time of the Continental Congress up through the late 1930s; ICPR planned to make this massive mountain of records machine-readable (Allen 1967, 5). ICPR also embraced an ambitious project to collect huge quantities of older U.S. election results; its role would be to standardize them and make them available for secondary analyses. This project, sponsored initially by the SSRC's Committee on Political Behavior and the American Historical Association, sought to collect information about U.S. elections as far back as 1824. This meant first locating the most complete and reliable county-level election results across the U.S. (a herculean task overseen by Walter Dean Burnham over several years in the early 1960s) and then, at ICPR, "cleaning" or standardizing these results to make them machine-readable. NSF sponsored the ICPR phase of this project beginning in 1963 with a series of substantial grants (NSF Grant Proposal

1964, 2-5; Austin 2011).<sup>15</sup> As Rokkan (1976) suggests in his discussion of Norwegian institutions, the collection and standardization of these kinds of public records about political decision-making would, in other contexts, unambiguously be a state agency's job. That it was open to a group of academic institutions like ICPR to undertake this work in the U.S., however, does not make the project less state-like.

ICPR also went about acquiring large quantities of survey research and public opinion data already in 1962, its first year of operation. It received the data and supporting materials for important studies from other survey research centers in the U.S., like Columbia's Bureau of Applied Social Research (BASR) and Chicago's National Opinion Research Center (NORC) (Miller 1962a). Each of these centers had contracts to do public opinion studies for governmental and corporate clients – as did Michigan's Survey Research Center (SRC) itself. So while ICPR did not acquire polling data from Roper or Gallup, it acquired a substantial amount of such data from other survey research centers which, like the SRC, had one foot in the university world and another in the world of commercial polling. For example, in the early 1960s, officials at NORC offered to withdraw multiple years of public opinion studies they had donated to the Roper Center if ICPR wanted them; NORC had done these studies of American public opinion about U.S. foreign policy for the U.S. State Department for over a decade in the 1940s and 50s. Miller balked at the fee Roper would probably charge for releasing these studies – but not at the idea of acquiring these studies themselves (Miller 1962b).<sup>16</sup> Though there were some suggestions that different repositories across the U.S. confine themselves to particular areas in which they would seek to acquire data, ICPR did not impose such limitations on itself. It welcomed data produced by public entities,

---

<sup>15</sup> Austin (2011) offers an excellent account of ICPR's early history. Austin joined the staff of ICPR to assist specifically with its work on this massive quantity of election data.

<sup>16</sup> Converse (1987, 321-323) recounts the controversy that erupted in the late 1950s when the State Department's long-standing "confidential" contract with NORC became public.

survey researchers and pollsters focused not only on the U.S., but also on U.S. foreign relations and public opinion in other countries.<sup>17</sup>

I do not, however, want to overstate the difference between these two broad classes of data ICPR sought to acquire. As I noted above, collecting, standardizing and distributing machine-readable data on votes taken in the U.S. Congress is a state-like project; but so too, I believe, is the collecting, standardizing and distributing of public opinion data. Though no sub-division of any federal government agency charged with doing public opinion research lasted much past World War II in the U.S., the strong ties between the federal government, academic survey researchers and commercial pollsters endured (Simpson 1994, 54-57). And thanks to those ties, public opinion researchers often functioned, in Christopher Simpson's phrase, as "outposts of the government," doing surveys for a wide range of governmental entities: the Office of Naval Research, the Air Force, the State Department, the U.S. Information Agency, etc.<sup>18</sup> So even though ICPR did not restrict its acquisitions to information first produced by public entities, much of the data it acquired in its first decade bore the state's stamp. Either it had been produced by officials working for local, state or federal entities or its production had been commissioned by them.

Only in the mid-1970s did a senior ICPR staff member question the unrelenting pace and scope of its acquisitions policy. Richard Hofferbert, who was about to leave his position as the director of ICPR for a position at another university, raised what he expected would be a jarring question: "Why do we want more data?" There had been so many acquisitions during his tenure as director, Hofferbert explained, that he couldn't possibly keep track of all of them or offer informed responses to donors who contacted him. What ICPR needed was a more deliberately constructed acquisitions policy, one that took into account what it already had, which portions of its collection were used and cited the most widely and so on (Hofferbert 1975, 2-3). Hofferbert made these comments as he was preparing to leave ICPR; I do

---

<sup>17</sup> For the sake of comparison, see Hastings (1963) for an overview of Roper's broad holdings.

<sup>18</sup> The phrase I've cited is the title of Chapter 5 of Simpson (1994).

not know whether his successors heeded them. Still, these comments provide more evidence that in its first decade, ICPR's acquisitions policy had been wide-ranging and aggressive. Not only had ICPR been pursuing particular batches of data; its default policy seems to have been to accept nearly any data that was offered to it.

### *Intellectual Aspirations that Fueled the Boom*

Up until this point, I have focused primarily on the competition between ICPR and the Roper Center in the 1960s. For as real and bitter as this rivalry was, I do not want to suggest that it was fueled solely by professional or institution-building motives. Had those who participated in it not *also* been driven by powerful intellectual aspirations, this rivalry and indeed the whole data center boom would not have been as intense as it was. What were some of those aspirations?

Those who devoted their professional lives to building data centers believed they could generate far deeper and broader knowledge about society and politics than what early 20<sup>th</sup> century social science had been able to muster – provided these data centers were well-stocked, expertly staffed and widely used. Data sets housed in data centers where they could be analyzed using the latest computing technology held out the promise that social scientists might now be able to make general or even universal claims. In broad strokes, such intellectual aspirations predated ICPR by over a decade. For instance, in 1949, the developer of the Human Relations Area Files claimed that an analysis of the collection's vast array of data about interpersonal relations in hundreds of cultures across the world could “prove” that the family – perhaps even the nuclear family – was a universal feature of human culture (Scheuch 2003, 387-388). Similarly, in the late 1950s, the anthropologist who promoted a microcard storage system for his and other anthropologists' research records believed that his “database of dreams” was uniquely poised to offer a “global understanding of man” (Lemov 2015, 212). Grandiose as these aspirations may now seem, it is hard to imagine people dedicating themselves to the daily drudgery that it took to build and promote the infrastructure of social scientific data without the impetus of such heady visions.

When people at ICPR expressed their intellectual hopes for what the institution they were building might make possible, they too argued that the depth and breadth of computer-aided analyses of vast quantities of data promised to reveal general, fundamental patterns beyond the reach of human analysts on their own. Focusing first on the potential of their recently acquired comprehensive collection of county-level electoral results from 1824 to the present, they argued the value of turning these older results into machine-readable data and including them in their holdings went far beyond the merely additive. Combined with the more recent electoral data ICPR already had, “each new data collection represents an increment that is exponentially more useful and valuable” (Grant Proposal 1964, 2). The hope was that computer-aided “longitudinal” analyses of this huge mass of electoral data might yield general insights into long-term patterns of political behavior – analyses not even the largest, most well-informed team of researchers could perform on their own.

Casting a deeper, wider net into the political past was not the only thing people at ICPR hoped their new data center would do. They also imagined it would give their research an expansive, cross-cultural breadth as well. A scholar with access to a large number of data sets from all over the world could “check out his theoretical surmises over much more extended areas” rather than being forced to limit himself to a “case study” (Grant Proposal 1967, 81). And once different researchers could work with the same data sets simultaneously, this too promised to bring about “a major innovation in the social organization of research” (Grant Proposal 1967, 82). Some or all of these changes may have been what ICPR’s Technical Services Director, Ralph Bisco, had in mind when he imagined “the revolutionary potential of large, integrated collections of social science data” “maximally exploit[ed]” by powerful new computers (1967, 71).

But such anticipations of the new vistas of social and political knowledge computer-aided analyses might open up gave way easily to far bleaker views. As Joy Rohde (forthcoming) argues, the development of computer systems to predict international crises that began in the late 1960s was driven by a profound “epistemic insecurity.” The flipside of a sunny, confident optimism about all the computer-aided human mind can know is a crushing awareness of the slowness, partiality and incorrigible flaws of

human cognition. Though some of the early “the sky’s the limit” optimism about computer-aided social science has surely faded, the conviction that the most reliable social science must be grounded in data remains. Similarly, though social scientific academic data centers and all the services they provide lost their futuristic gleam decades ago, they have become unglamorous but indispensable institutions in the infrastructure of contemporary social science. Their enduring position in that infrastructure over the last nearly sixty years has had consequences far outside their bounds for how social and political inquiry is taught and practiced.

## References

### Archival Sources

#### ICPSR Records

Allen, Howard. 1967. Memo from Director of Data Recovery, Subject: "Progress Report on Historical Data Projects," May 30. Folder: Historical Archive, 1967. Box 14.

Converse, Philip E. 1962a. Letter to Karl Deutsch, Member of Committee of Eight. October 26. Folder: Committee of Eight, Correspondence and Minutes, 1962. Box 1.

\_\_\_\_\_. 1962b. Memo to Committee of 8, "Background, Issues & Tentative Agenda for Meeting, December 1-2," November 8. Folder: Committee of Eight, Correspondence and Minutes, 1962. Box 1.

\_\_\_\_\_. 1963a. Proposal, "Toward the Development of an Inter-University Data Archive System for Comparative Research," labeled "First Draft, January, 1963," hand-dated 1.31.63. Folder: Organization and Administration, Committee of Eight, Proposals, 1963. Box 1.

\_\_\_\_\_. 1963b. Letter to Charles Glock, Survey Research Center, U.C. Berkeley. May 2. Folder: Organization and Administration, Committee of Eight, Correspondence and Minutes, 1963. Box 1.

\_\_\_\_\_. 1963c. Memo to Committee of 8, June 28. Folder: Organization and Administration, Committee of Eight, Correspondence and Minutes, 1963. Box 1.

Grant Proposal. 1964. No Author. "A Proposal for a Study of the National Electorate, 1964." May 13. Box 16.

Grant Proposal. 1967. No Author. "To Ford Foundation for Funds to Support an Expansion of Archival Resources." Box 16.

Hofferbert, Richard. 1975. Memo to Jerome M. Clubb, "Supervision of Data Acquisition." Folder: Correspondence and Memoranda folder, 1964, 1966, 1974-75. Box 14.

\_\_\_\_\_. N.d. "Archiving of Machine-Readable Policy Data: Activities of the ICPR." Folder: Resource Development and Services, Archival Activities, Reports and Reviews, 1962-1978. Box 15.

Lane, Robert. 1964. Letter to Miller, November 7. Folder: Correspondence To/From ICPSR Executive Directors, 1964-2001, 1 of 3, Box 3.

Lowi, Theodore. 1964. Memo to Miller and John Wahlke. "Anti- 'Inter-University Cooperation'," November 2. Folder: Correspondence To/From ICPSR Executive Directors, 1964-2001, 1 of 3, Box 3.

Miller, Warren E. 1962a. Memo to ICPR Council, "Estimated cost of establishment of data repository," October 18. Folder: Resource Development and Services, Archival Activities: Formation of a Data Repository, 1962, Box 14.

\_\_\_\_\_. 1962b. Memo to ICPR Council, "Conversations with Peter Rossi and Jack Feldman of NORC, October 4, 1962" n.d. Folder: Resource Development and Services, Archival Activities: Formation of a Data Repository, 1962, Box 14.

\_\_\_\_\_. 1964a. Memo to Official Representatives, ICPR, re. "Inter-archival cooperation," September 1964. Folder: Correspondence To/From ICPSR Executive Directors, 1964-2001, 1 of 3, Box 3.

\_\_\_\_\_. 1964b. "Proposed Principles for Inter-Archive Cooperation," n.d.; attached to Miller, 1964a. Folder: Correspondence To/From ICPSR Executive Directors, 1964-2001, 1 of 3, Box 3.

National Science Foundation Grant Proposal (no author). 1964. A Proposal for Funds to Support the Addition of Data to the Inter-University Consortium for Political Research Data Repository, February. Box 16.

Wahlke, John. 1964. Council, ICPR, to Official Representative, ICPR; Subject: "Proposed Principles for Inter-Archival Cooperation," Draft, n.d. Folder: Correspondence To/From ICPSR Executive Directors, 1964-2001, 1 of 3, Box 3.

Miller, Warren E. 1997. Oral history interview. Oral history interview. (Erik Austin, interviewer). July 20. VHS tape and typed transcript. Institute for Social Research (ISR) Oral History project.

Miller, Warren E. Personal Papers.

1966. Proposal Rating Sheet, Proposal S 70163-P (An African Social Science Data Archive). Folder: Correspondence, Topical - Proposal Reviews, NSF: Miscellaneous, 1962-67. Box 3.

### Published and Online Works

Austin, Erik. 2011. ICPSR: The Founding and Early Years. Accessed at <http://www.icpsr.umich.edu/icpsrweb/content/membership/history/early-years.html>.

Bisco, Ralph L. 1966. Social Science Data Archives: A Review of Developments. *The American Political Science Review* 60 (1): 93-109.

\_\_\_\_\_. 1967. Social Science Data Archives: Progress and Prospects. *Social Science Information* 6 (1): 39-74.

Converse, Jean M. 1987. *Survey Research in the United States: Roots and Emergence, 1890-1960*. Berkeley: University of California Press.

Farber, Karen A. and Carl Beck. 1971. Council of Social Science Data Archives (CSSDA). *Encyclopedia of Library and Information Science*.

Garfinkel, Simson. 2000. *Database Nation. The Death of Privacy in the 21<sup>st</sup> Century*. Sebastopol, CA: O'Reilly and Associates.

Hastings, Philip K. 1963. The Roper Center: An International Archive of Sample Survey Data. *Public Opinion Quarterly* 27 (4): 590-598.

- Hauptmann, Emily. 2017. Why They Shared. Recovering Early Arguments for Sharing Social Scientific Data. Revised version of paper presented at the 2016 IPSA World Congress, Poznań, Poland. Under review. Available on request.
- Kraus, Rebecca S. 2011. Statistical Déjà Vu: The National Data Center Proposal of 1965 and its Descendants. Paper presented at the Joint Statistical Meetings, Miami Beach, FL., August 1.
- Lemov, Rebecca. 2015. *Database of Dreams: The Lost Quest to Catalog Humanity*. New Haven, CT: Yale University Press.
- Lucci, York and Stein Rokkan, with Eric Meyerhoff. 1957. *A Library Center of Survey Research Data: A Report of an Inquiry and a Proposal*. New York: Columbia University School of Library Service.
- Miller, Arthur R. 1971. *The Assault on Privacy: Computers, Data Banks, and Dossiers*. Ann Arbor, MI: University of Michigan Press.
- Rohde, Joy. Forthcoming. Pax Technologica. Computers, International Affairs and Human Reason in the Age of American Hegemony. *Isis*.
- Rokkan, Stein. 1976. Data Services in Western Europe. Reflections on Variations in the Conditions of Academic Institution-Building. *American Behavioral Scientist* 19 (4): 443-454.
- Scheuch, Erwin K. 2003. History and Visions in the Development of Data Services for the Social Sciences. *International Social Science Journal* 177: 385-399.
- Simpson, Christopher. 1994. *Science of Coercion. Communication Research and Psychological Warfare, 1945-1960*. Oxford: Oxford University Press.
- Truman, David B. 1991. Oral history interview. In Michael A. Baer, Malcolm E. Jewell and Lee Sigelman, eds. *Political Science in America: Oral Histories of a Discipline*. Lexington, KY.: University of Kentucky Press.